INDIVIDUAL DEGREES OF FREEDOM FOR NON-ADDITIVITY IN TWO- AND HIGHER-WAY DESIGNS

BU-136-M Janet C. Cassady October, 1961

Abstract

In this paper is presented a method by which the error sum of squares in two- and higher-way designs may be partitioned into single degree of freedom sums of squares such that one of the sums of squares so isolated for the two-way design is Tukey's [1949] sum of squares for non-additivity. Also presented is a proof that each of these sums of squares, under the assumptions underlying the analysis of variance, has a σ^2 chi-square distribution with one degree of freedom, independent of the remaining sums of squares.

INDIVIDUAL DEGREES OF FREEDOM FOR NON-ADDITIVITY
IN TWO- AND HIGHER-WAY DESIGNS

BU-136-M Janet C. Cassady October, 1961

Introduction

The one degree of freedom sum of squares for non-additivity introduced by Tukey [1949] is a most useful statistical tool. One of the basic assumptions for the use of the analysis of variance is that of additivity of effects [Eisenhart, 1947]. In the two-way design, for example, we assume that the X_{ij} have a joint multivariate normal distribution, with mean vector $\underline{\mu+\alpha_i+\beta_j}$ and covariance matrix of σ^2 times the identity matrix $\underline{I}$.

Testing the validity of the assumption of additivity of effects, then, is an important consideration. A procedure to do this has been described by Tukey [1949] for an r x c two-way classification with one observation per cell and for other situations and for more than one observation per subclass [see Abraham, 1960; Elston, 1959, 1961; Federer, 1955, 1959; Hamaker, 1955; Harter and Lum, 1957; Moore and Tukey, 1954; Snedecor, 1956; Tukey, 1955; Ward and Dick, 1952]. The procedure Tukey [1949] described partitions the interaction sum of squares in an r x c classification into two parts: one is the sum of squares for non-additivity, having under the null hypothesis a σ^2 chi-square distribution with one degree of freedom; the other is the remainder of the interaction sum of squares, having under the null hypothesis a σ^2 chi-square distribution with rc-c-r degrees of freedom. The ratio of the first sum of squares to the mean square of the residual, then, under the null hypothesis, is said to have an F-distribution with 1 and rc-c-r degrees of freedom [Tukey, 1949].

It is the purpose of this paper, first, to propose a method by which the error sum of squares in two- and higher-way designs may be partitioned into single degree of freedom sums of squares such that one of the sums of squares so isolated for the two-way design is Tukey's [1949] sum of squares for non-additivity; and second, to present a proof that each of these sums of squares, under the null hypothesis, has a σ^2 chi-square distribution with one degree of freedom, independent of the remaining sums of squares.

<u>Individual Degrees of Freedom in an r x c Classification</u>

In the two-factor treatment design with known levels in both directions, it has been common practice to use orthogonal polynomials to construct orthogonal contrasts among levels of the treatment factors. Thus, the sum of squares due to treatment A at r levels may be partitioned into r-1 sums of squares: first, A_L, the sum of squares due to linear regression on the levels of A; second, A_Q, the additional sum of squares due to a quadratic regression on the levels of A; third, A_C, the additional sum of squares due to a cubic regression; and so forth. The B sum of squares may be partitioned in a similar manner.

In such a case, the interaction sum of squares may be partitioned into (r-1)(c-1) independent sums of squares, using as coefficients the Kronecker product of the two matrices of coefficients for the treatment factors.

When the treatment levels are equally spaced, the computations are simple. However, Robson [1959] has suggested a relatively simple method for constructing orthogonal polynomials when the independent variable is unequally spaced. This method becomes even simpler to use when it has been programmed for a modern electronic computer.

Now, let us consider the case at hand, in which we do not know the level of the factors. If we could assume that the factors are unequally spaced according to the observed mean levels for the r rows $\bar{x}_{1.}$, $\bar{x}_{2.}$, ..., $\bar{x}_{r.}$ and for the columns $\bar{x}_{.1}$, $\bar{x}_{.2}$, ..., $\bar{x}_{.c}$, we might use Robson's [1959] method for constructing orthogonal contrasts when the independent variable is unequally spaced. Partitioning the interaction sum of squares in this fashion (using as coefficients the Kronecker product of the two matrices of coefficients computed assuming the rows and columns are spaced according to the observed means), we find that the one degree of freedom sum of squares computed corresponding to the $A_L B_L$ comparison would be Tukey's one degree of freedom sum of squares for non-additivity. Also, this is the sum of squares due to linear regression of the error estimate on the product of the corresponding row and column effect estimates [Harter and Lum, 1957].

In addition, Robson and Atkinson [1960] have demonstrated that the among regression coefficients sum of squares in a one-way analysis of covariance also may be partitioned into individual degree of freedom sums of squares.

<u>Distribution of the Ratio of Tukey's One Degree of Freedom for Non-Additivity Sum of Squares to the Residual Mean Square</u>

First we shall present a proof that Tukey's sum of squares for non-additivity has, under the basic assumptions of the analysis of variance [Eisenhart, 1947], a σ^2 chi-square distribution. Tukey [1949] stated that this is so but did not present an analytical proof. This is an important preliminary in our further partitioning of the interaction sum of squares.

We assume, as stated before, that the X_{ij} have a joint multivariate normal distribution, $N(\underline{\mu+\alpha_i+\beta_j}, \sigma^2[I])$.

We define our parameter estimates as follows:

$$\hat{\mu} = \bar{x}_{..} = \frac{1}{rc}\sum_{i,j}^{r,c} X_{ij} \text{ , } i = 1, \cdots, r \text{ ; } j = 1, \cdots, c$$

$$\hat{\alpha}_i = \bar{x}_{i.} - \bar{x}_{..} = \frac{1}{c}\sum_{j}^{c} X_{ij} - \bar{x}_{..}$$

$$\hat{\beta}_j = \bar{x}_{.j} - \bar{x}_{..} = \frac{1}{r}\sum_{i}^{r} X_{ij} - \bar{x}_{..}$$

$$\hat{\epsilon}_{ij} = X_{ij} - \bar{x}_{..} - \hat{\alpha}_i - \hat{\beta}_j = X_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..}$$

The variables $\hat{\mu}$, $\hat{\alpha}_1, \cdots, \hat{\alpha}_r, \hat{\beta}_1, \cdots, \hat{\beta}_c, \hat{\epsilon}_{11}, \cdots, \hat{\epsilon}_{rc}$ are linear combinations of variables with a joint multivariate normal distribution. These parameter estimates then have a joint multivariate normal distribution with a singular covariance matrix, since there are $1 + r + c + rc$ parameter estimates derived from only rc observed X_{ij}'s.

Let us restrict our consideration to a nonsingular transformation of the observations into a new set of rc random variables (a subset of the estimated parameters).

$$\underline{Z} = \underline{A} \cdot \begin{bmatrix} X_{11} \\ X_{12} \\ \cdots \\ X_{1c} \\ X_{21} \\ \cdots \\ X_{2c} \\ \cdots \\ X_{r1} \\ X_{r2} \\ \cdots \\ X_{rc} \end{bmatrix} = \begin{bmatrix} \bar{x} \\ \hat{\alpha}_1 \\ \cdots \\ \hat{\alpha}_{r-1} \\ \hat{\beta}_1 \\ \cdots \\ \hat{\beta}_{c-1} \\ \hat{\epsilon}_{11} \\ \cdots \\ \hat{\epsilon}_{1,\ c-1} \\ \hat{\epsilon}_{21} \\ \cdots \\ \hat{\epsilon}_{r-1,\ c-1} \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a\\a\\a\\a\\a\\a\\a \end{matrix}}\right\} \underline{Z}^{(1)} \\ \left.\vphantom{\begin{matrix} a\\a\\a\\a\\a\\a \end{matrix}}\right\} \underline{Z}^{(2)} \end{matrix}$$

The parameter estimates excluded from the right-hand column vector are all functions of included parameter estimates.

That is, since $\sum_i^r \hat{\alpha}_i = 0$,

$$\hat{\alpha}_r = -\sum_i^{r-1} \hat{\alpha}_i \ .$$

Similarly, $\hat{\beta}_c = -\sum_i^{c-1} \hat{\beta}_j$

$$\hat{\epsilon}_{ic} = -\sum_j^{c-1} \hat{\epsilon}_{ij} \qquad i \neq r$$

$$\hat{\epsilon}_{rj} = -\sum_i^{r-1} \hat{\epsilon}_{ij} \qquad j \neq c$$

$$\hat{\epsilon}_{rc} = \sum_i^{r-1} \sum_j^{c-1} \hat{\epsilon}_{ij} \ .$$

The transformation matrix $\underline{A}$ appears as follows:

$$
\begin{array}{rc}
 & \begin{array}{cccccc} \phantom{1-\frac{1}{r}-\frac{1}{c}+\frac{1}{rc}} & \phantom{1-\frac{1}{r}-\frac{1}{c}+\frac{1}{rc}} & & \text{Column} & \phantom{1-\frac{1}{r}-\frac{1}{c}+\frac{1}{rc}} & \\ & & & (c+1) & & \end{array} \\
\begin{array}{r}
\\[6pt] \\[6pt] \\[6pt] \\[6pt] \text{Row } (r+1) \\[6pt] \\[6pt] \\[6pt] \text{Row } (r+c) \\[6pt] \\[6pt] \\[6pt] \text{Row } (r+2c-1) \\[6pt] \\[6pt] \\
\end{array}
&
\left[
\begin{array}{cccccc}
\frac{1}{rc} & \frac{1}{rc} & \cdots & \frac{1}{rc} & \frac{1}{rc} & \cdots \\[6pt]
\frac{1}{c}-\frac{1}{rc} & \frac{1}{c}-\frac{1}{rc} & \cdots & -\frac{1}{rc} & -\frac{1}{rc} & \cdots \\[6pt]
-\frac{1}{rc} & -\frac{1}{rc} & \cdots & \frac{1}{c}-\frac{1}{rc} & \frac{1}{c}-\frac{1}{rc} & \cdots \\[6pt]
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\[6pt]
\frac{1}{r}-\frac{1}{rc} & -\frac{1}{rc} & \cdots & \frac{1}{r}-\frac{1}{rc} & -\frac{1}{rc} & \cdots \\[6pt]
-\frac{1}{rc} & \frac{1}{r}-\frac{1}{rc} & \cdots & -\frac{1}{rc} & \frac{1}{r}-\frac{1}{rc} & \cdots \\[6pt]
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\[6pt]
1-\frac{1}{r}-\frac{1}{c}+\frac{1}{rc} & -\frac{1}{c}+\frac{1}{rc} & \cdots & -\frac{1}{r}+\frac{1}{rc} & +\frac{1}{rc} & \cdots \\[6pt]
-\frac{1}{c}+\frac{1}{rc} & 1-\frac{1}{r}-\frac{1}{c}+\frac{1}{rc} & \cdots & \frac{1}{rc} & -\frac{1}{r}+\frac{1}{rc} & \cdots \\[6pt]
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\[6pt]
-\frac{1}{r}+\frac{1}{rc} & +\frac{1}{rc} & \cdots & 1-\frac{1}{r}-\frac{1}{c}+\frac{1}{rc} & -\frac{1}{c}+\frac{1}{rc} & \cdots \\[6pt]
\frac{1}{rc} & -\frac{1}{r}+\frac{1}{rc} & \cdots & -\frac{1}{c}+\frac{1}{rc} & 1-\frac{1}{r}-\frac{1}{c}+\frac{1}{rc} & \cdots \\[6pt]
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{array}
\right]
\end{array}
$$

Now, since $\underline{X}$ is distributed according to $N(\underline{\mu+\alpha_i+\beta_j}\ , \sigma^2[I])$, $\underline{Z} = \underline{AX}$ is distributed [Anderson, 1958] according to $N(\underline{A}\cdot\mu+\alpha_i+\beta_j\ , \sigma^2\underline{AIA}')$.

$$\underline{A}\cdot\underline{\mu+\alpha_i+\beta_j} = \underline{A}\cdot\begin{bmatrix} \mu+\alpha_1+\beta_1 \\ \mu+\alpha_1+\beta_2 \\ \cdots \\ \\ \\ \\ \\ \\ \\ \mu+\alpha_r+\beta_c \end{bmatrix} = \begin{bmatrix} \left.\begin{matrix} \mu+\frac{1}{r}\Sigma\alpha_i+\frac{1}{c}\Sigma\beta_j \\ \alpha_1-\frac{1}{r}\Sigma\alpha_i \\ \cdots \\ \alpha_{r-1}-\frac{1}{r}\Sigma\alpha_i \\ \beta_1-\frac{1}{c}\Sigma\beta_j \\ \cdots \end{matrix}\right\} \underline{\mu}^{(1)} \\ \left.\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}\right\} \underline{\mu}^{(2)} \end{bmatrix}$$

$\sigma^2\underline{AA}'$ then is the covariance matrix, which is in the form of an array of several independent covariance matrices,

$$\sigma^2\underline{AA}' = \begin{bmatrix} \mathrm{Var}(\bar{x}) & 0 & 0 & 0 \\ 0 & \mathrm{Cov}(\underline{\hat{\alpha}}_i,\underline{\hat{\alpha}}_{i'}) & 0 & 0 \\ 0 & 0 & \mathrm{Cov}(\underline{\hat{\beta}}_j,\underline{\hat{\beta}}_{j'}) & 0 \\ 0 & 0 & 0 & \mathrm{Cov}(\underline{\hat{\epsilon}}_{ij},\underline{\hat{\epsilon}}_{i'j'}) \end{bmatrix}$$

The individual elements of the matrix obtained from the indicated matrix multiplication are as itemized below; first the computation is shown; then the elements are arrayed in the several covariance matrices. Thus,

$$\mathrm{Var}(\bar{x}) = \sigma^2 \sum^{rc}\left(\frac{1}{rc}\right)^2 = \frac{\sigma^2}{rc}$$

$$\mathrm{Var}(\hat{\alpha}_i) = \sigma^2\left[\sum^{c}\left(\frac{1}{c}-\frac{1}{rc}\right)^2 + \sum^{rc-c}\left(-\frac{1}{rc}\right)^2\right] = \frac{r-1}{rc}\,\sigma^2$$

$$\mathrm{Cov}(\hat{\alpha}_i,\hat{\alpha}_{i'}) = \sigma^2 \left[\sum^{2c} \left(-\frac{1}{rc}\right)\left(\frac{1}{c} - \frac{1}{rc}\right) + \sum^{(r-2)c} \left(-\frac{1}{rc}\right)^2 \right] = \sigma^2 \left(-\frac{1}{rc}\right)$$

$$\mathrm{Var}(\hat{\beta}_j) = \sigma^2 \left[\sum^{r} \left(\frac{1}{r} - \frac{1}{rc}\right)^2 + \sum^{rc-r} \left(-\frac{1}{rc}\right)^2 \right] = \frac{(c-1)}{rc}\sigma^2$$

$$\mathrm{Cov}(\hat{\beta}_j,\hat{\beta}_{j'}) = \sigma^2 \left[\sum^{2r} \left(-\frac{1}{rc}\right)\left(\frac{1}{r} - \frac{1}{rc}\right) + \sum^{(c-2)r} \left(-\frac{1}{rc}\right)^2 \right] = \frac{c-1}{rc}\sigma^2$$

$$\begin{aligned}
\mathrm{Cov}(\hat{\alpha}_i,\hat{\beta}_j) &= \sigma^2 \left[\left(\frac{1}{r} - \frac{1}{rc}\right)\left(\frac{1}{c} - \frac{1}{rc}\right) + \sum^{r-1} \left(-\frac{1}{rc}\right)\left(\frac{1}{r} - \frac{1}{rc}\right) \right. \\
&\quad \left. + \sum^{c-1} \left(-\frac{1}{rc}\right)\left(\frac{1}{c} - \frac{1}{rc}\right) + \sum^{rc-r-c-1} \left(-\frac{1}{rc}\right)^2 \right] = 0
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(\hat{\beta}_j,\hat{\epsilon}_{ij}) &= \sigma^2 \left[\left(\frac{1}{r} - \frac{1}{rc}\right)\left(1 - \frac{1}{r} - \frac{1}{c} + \frac{1}{rc}\right) + \sum^{r-1} \left(\frac{1}{r} - \frac{1}{rc}\right)\left(-\frac{1}{r} + \frac{1}{rc}\right) \right. \\
&\quad \left. + \sum^{c-1} \left(-\frac{1}{c} + \frac{1}{rc}\right)\left(-\frac{1}{rc}\right) + \sum^{rc-r-c+1} \left(-\frac{1}{rc}\right)\left(+\frac{1}{rc}\right) \right] = 0
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(\hat{\epsilon}_{ij}) &= \sigma^2 \left[\left(1 - \frac{1}{r} - \frac{1}{c} + \frac{1}{rc}\right)^2 + \sum^{r-1} \left(-\frac{1}{r} + \frac{1}{rc}\right)^2 + \sum^{c-1} \left(-\frac{1}{c} + \frac{1}{rc}\right)^2 \right. \\
&\quad \left. + \sum^{rc-r-c+1} \left(\frac{1}{rc}\right)^2 \right] = \frac{(r-1)(c-1)}{rc}\sigma^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(\hat{\epsilon}_{ij},\hat{\epsilon}_{i'j}) &= \sigma^2 \left[2\left(1 - \frac{1}{r} - \frac{1}{c} + \frac{1}{rc}\right)\left(-\frac{1}{c} + \frac{1}{rc}\right) + (c-2)\left(-\frac{1}{c} + \frac{1}{rc}\right)^2 \right. \\
&\quad \left. + 2(r-1)\left(-\frac{1}{r} + \frac{1}{rc}\right)\left(\frac{1}{rc}\right) + (rc-c-2r+2)\left(\frac{1}{rc}\right)^2 \right] = -\frac{(r-1)}{rc}\sigma^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(\hat{\epsilon}_{ij},\hat{\epsilon}_{i'j'}) &= \sigma^2 \left[2\left(1 - \frac{1}{r} - \frac{1}{c} + \frac{1}{rc}\right)\left(\frac{1}{rc}\right) + 2(c-2)\left(-\frac{1}{c} + \frac{1}{rc}\right)\left(\frac{1}{rc}\right) \right. \\
&\quad + 2(c-2)\left(-\frac{1}{r} + \frac{1}{rc}\right)\left(\frac{1}{rc}\right) + 2\left(-\frac{1}{c} + \frac{1}{rc}\right)\left(-\frac{1}{r} + \frac{1}{rc}\right) \\
&\quad \left. + (rc-2c-2r+4)\left(\frac{1}{r^2c^2}\right) \right] = \frac{1}{rc}\sigma^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(\hat{\alpha}_i,\hat{\epsilon}_{ij}) &= \sigma^2 \left[\left(\frac{1}{c} - \frac{1}{rc}\right)\left(1 - \frac{1}{r} - \frac{1}{c} + \frac{1}{rc}\right) + \sum^{c-1} \left(\frac{1}{c} - \frac{1}{rc}\right)\left(-\frac{1}{c} + \frac{1}{rc}\right) \right. \\
&\quad \left. + \sum^{r-1} \left(-\frac{1}{r} + \frac{1}{rc}\right)\left(-\frac{1}{rc}\right) + \sum^{rc-r-c+1} \left(-\frac{1}{rc}\right)\left(\frac{1}{rc}\right) \right] = 0
\end{aligned}$$

$$\mathrm{Cov}(\hat{\beta}_j, \hat{\epsilon}_{ij'}) = \sigma^2 \Big[\left(\frac{1}{r} - \frac{1}{rc}\right)\left(-\frac{1}{c} + \frac{1}{rc}\right) + \left(1 - \frac{1}{r} - \frac{1}{c} + \frac{1}{rc}\right)\left(-\frac{1}{rc}\right)$$

$$+ \sum^{r-1}\left(\frac{1}{rc}\right)\left(\frac{1}{r} - \frac{1}{rc}\right) + \sum^{r-1}\left(-\frac{1}{rc}\right)\left(-\frac{1}{r} + \frac{1}{rc}\right)$$

$$+ \sum^{c-2}\left(-\frac{1}{c} + \frac{1}{rc}\right)\left(-\frac{1}{rc}\right) + \sum^{rc-2r-c+2}\left(\frac{1}{rc}\right)\left(-\frac{1}{rc}\right)\Big] = 0$$

$$\mathrm{Cov}(\hat{\alpha}_i, \hat{\epsilon}_{i'j}) = \sigma^2 \Big[\left(\frac{1}{c} - \frac{1}{rc}\right)\left(-\frac{1}{r} + \frac{1}{rc}\right) + \left(1 - \frac{1}{r} - \frac{1}{c} + \frac{1}{rc}\right)\left(-\frac{1}{rc}\right)$$

$$+ \sum^{c-1}\left(\frac{1}{rc}\right)\left(\frac{1}{c} - \frac{1}{rc}\right) + \sum^{c-1}\left(-\frac{1}{rc}\right)\left(-\frac{1}{c} + \frac{1}{rc}\right)$$

$$+ \sum^{r-2}\left(-\frac{1}{r} + \frac{1}{rc}\right)\left(-\frac{1}{rc}\right) + \sum^{rc-2c-r+2}\left(\frac{1}{rc}\right)\left(-\frac{1}{rc}\right)\Big] = 0$$

$$\mathrm{Cov}(\hat{\epsilon}_{ij}, \hat{\epsilon}_{ij'}) = \sigma^2 \Big[2\left(1 - \frac{1}{r} - \frac{1}{c} + \frac{1}{rc}\right)\left(-\frac{1}{r} + \frac{1}{rc}\right) + (r-2)\left(-\frac{1}{r} + \frac{1}{rc}\right)^2$$

$$+ 2(c-1)\left(-\frac{1}{c} + \frac{1}{rc}\right)\left(\frac{1}{rc}\right) + (rc-r-2c+2)\left(\frac{1}{rc}\right)^2\Big] = -\frac{(c-1)}{rc}\sigma^2$$

Thus, the covariance matrices appear as follows:

$$\begin{bmatrix} \mathrm{Var}(\hat{\alpha}_1) & \mathrm{Cov}(\hat{\alpha}_1,\hat{\alpha}_2) & \cdots & \mathrm{Cov}(\hat{\alpha}_1,\hat{\alpha}_{r-1}) \\ \mathrm{Cov}(\hat{\alpha}_1,\hat{\alpha}_2) & \mathrm{Var}(\hat{\alpha}_2) & \cdots & \mathrm{Cov}(\hat{\alpha}_2,\hat{\alpha}_{r-1}) \\ \mathrm{Cov}(\hat{\alpha}_1,\hat{\alpha}_3) & \mathrm{Cov}(\hat{\alpha}_2,\hat{\alpha}_3) & \cdots & \mathrm{Cov}(\hat{\alpha}_3,\hat{\alpha}_{r-1}) \\ \cdots & \cdots & \cdots & \cdots \\ \mathrm{Cov}(\hat{\alpha}_1,\hat{\alpha}_{r-1}) & \mathrm{Cov}(\hat{\alpha}_2,\hat{\alpha}_{r-1}) & \cdots & \mathrm{Var}(\hat{\alpha}_{r-1}) \end{bmatrix} = \frac{\sigma^2}{rc} \begin{bmatrix} r-1 & -1 & \cdots & -1 \\ -1 & r-1 & \cdots & -1 \\ -1 & -1 & \cdots & -1 \\ \cdots & \cdots & \cdots & \cdots \\ -1 & -1 & \cdots & r-1 \end{bmatrix}$$

$$\begin{bmatrix} \mathrm{Var}(\hat{\beta}_1) & \mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_2) & \cdots & \mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_{c-1}) \\ \mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_2) & \mathrm{Var}(\hat{\beta}_2) & \cdots & \mathrm{Cov}(\hat{\beta}_2,\hat{\beta}_{c-1}) \\ \cdots & \cdots & \cdots & \cdots \\ \mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_{c-1}) & \mathrm{Cov}(\hat{\beta}_2,\hat{\beta}_{c-1}) & \cdots & \mathrm{Var}(\hat{\beta}_{c-1}) \end{bmatrix} = \frac{\sigma^2}{rc} \begin{bmatrix} c-1 & -1 & \cdots & -1 \\ -1 & c-1 & \cdots & -1 \\ \cdots & \cdots & \cdots & \cdots \\ -1 & -1 & \cdots & c-1 \end{bmatrix}$$

$$\begin{bmatrix} \mathrm{Var}(\hat\epsilon_{11}) & \mathrm{Cov}(\hat\epsilon_{11},\hat\epsilon_{12}) & \cdots & \mathrm{Cov}(\hat\epsilon_{11},\hat\epsilon_{21}) & \mathrm{Cov}(\hat\epsilon_{11},\hat\epsilon_{22}) & \cdots \\ \mathrm{Cov}(\hat\epsilon_{11},\epsilon_{12}) & \mathrm{Var}(\hat\epsilon_{12}) & \cdots & \mathrm{Cov}(\hat\epsilon_{12},\hat\epsilon_{21}) & \mathrm{Cov}(\hat\epsilon_{12},\hat\epsilon_{22}) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathrm{Cov}(\hat\epsilon_{11},\hat\epsilon_{21}) & \mathrm{Cov}(\hat\epsilon_{12},\hat\epsilon_{21}) & \cdots & \mathrm{Var}(\hat\epsilon_{21}) & \mathrm{Cov}(\hat\epsilon_{21},\hat\epsilon_{22}) & \cdots \\ \mathrm{Cov}(\hat\epsilon_{11},\hat\epsilon_{22}) & \mathrm{Cov}(\hat\epsilon_{12},\hat\epsilon_{22}) & \cdots & \mathrm{Cov}(\hat\epsilon_{21},\hat\epsilon_{22}) & \mathrm{Var}(\hat\epsilon_{22}) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$= \frac{\sigma^2}{rc} \begin{bmatrix} (r-1)(c-1) & -(c-1) & \cdots & -(r-1) & 1 & \cdots \\ -(c-1) & (r-1)(c-1) & \cdots & 1 & -(r-1) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -(r-1) & 1 & \cdots & (r-1)(c-1) & -(c-1) & \cdots \\ 1 & -(r-1) & \cdots & -(c-1) & (r-1)(c-1) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

It is now apparent that the vector of $\hat\alpha_i$ estimates is independent of the vector of $\hat\beta_j$ estimates and that both vectors are independent of $\bar{x}$ and the vector of $\hat\epsilon_{ij}$ estimates, since a necessary and sufficient condition [Anderson, 1958] that one subset of the random variables of a joint normal distribution and the subset consisting of the remaining variables be independent is that each covariance of a variable from one set and a variable from the other set be 0.

Using the parameter estimates we have defined previously, we should like to prove $\hat\gamma_1 = \dfrac{\sum\limits_{i,j}^{rc} \hat\alpha_i \hat\beta_j \epsilon_{ij}}{\sqrt{\sum\limits_i^r \hat\alpha_i^2 \sum\limits_j^c \hat\beta_j^2}}$ is $N(0,\sigma^2)$. The square of this quantity is the sum of squares for Tukey's one degree of freedom.

Several well-known lemmas, e.g. Madow [1949], will be useful in the proof.

Lemma 1. $E(X) = E[E(X|Z)]$

Lemma 2. $Var(X) = E[Var(X|Z)] + Var[E(X|Z)]$

Lemma 3. $Cov(X,Y) = E[Cov(X,Y|Z)] + Cov[E(X|Z), E(Y|Z)]$

Using Lemma 1,

$$E\left[\frac{\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}}\right] = E\left\{E\left[\frac{\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}} \;\middle|\; \hat{\alpha}_1,\cdots,\hat{\alpha}_{r-1},\hat{\beta}_1,\cdots,\hat{\beta}_{c-1}\right]\right\}$$

Since the subsets $\hat{\epsilon}_{ij}$, $\hat{\alpha}_i$, and $\hat{\beta}_j$ are independent of each other, the above expression is equal to

$$E\left[\frac{\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j E(\hat{\epsilon}_{ij})}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}}\right] = 0$$

Thus, $E(\hat{\gamma}_1) = 0$.

Using Lemma 2, we may write

$$V\left[\frac{\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}}\right] = E\left\{V\left[\frac{\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}} \;\middle|\; \hat{\alpha}_1,\cdots,\hat{\alpha}_{r-1},\hat{\beta}_1,\cdots,\hat{\beta}_{c-1}\right]\right\}$$

$$+ V\left\{E\left[\frac{\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}} \;\middle|\; \hat{\alpha}_1,\cdots,\hat{\alpha}_{r-1},\hat{\beta}_1,\cdots,\hat{\beta}_{c-1}\right]\right\}$$

From the previous expectation, we know that the second term of this expression is zero. Therefore, the above expression becomes

$$E\left\{\frac{1}{\sum_i^r \hat{\alpha}_i^2 \sum_j^c \hat{\beta}_j^2}\left[\sum_{i,j}^{r,c} \hat{\alpha}_i^2 \hat{\beta}_j^2\, V(\hat{\epsilon}_{ij}) + \sum_{\substack{i,j,j' \\ j\neq j'}}^{r,c,c} \hat{\alpha}_i^2 \hat{\beta}_j \hat{\beta}_{j'}\, \mathrm{Cov}(\hat{\epsilon}_{ij},\hat{\epsilon}_{ij'})\right.\right.$$

$$+ \sum_{\substack{i,i',j \\ i\neq i'}}^{r,r,c} \hat{\alpha}_i \hat{\alpha}_{i'} \hat{\beta}_j^2\, \mathrm{Cov}(\hat{\epsilon}_{ij},\hat{\epsilon}_{i'j})$$

$$\left.\left. + \sum_{\substack{i,i',j,j' \\ i\neq i' \\ j\neq j'}}^{r,r,c,c} \hat{\alpha}_i \hat{\alpha}_{i'} \hat{\beta}_j \hat{\beta}_{j'}\, \mathrm{Cov}(\hat{\epsilon}_{ij},\hat{\epsilon}_{i',j'})\right]\right\}$$

$$= E\left\{\frac{1}{\sum_i^r \hat{\alpha}_i^2 \sum_j^c \hat{\beta}_j^2}\left[\sum_{i,j}^{r,c} \hat{\alpha}_i^2 \hat{\beta}_j^2\, \frac{(r-1)(c-1)}{rc}\sigma^2\right.\right.$$

$$+ \sum_{i,j}^{r,c} \hat{\alpha}_i \hat{\beta}_j (-\hat{\beta}_j)(-\frac{(c-1)}{rc}\sigma^2) + \sum_{i,j}^{rc} \hat{\alpha}_i (-\hat{\alpha}_i) \hat{\beta}_j^2 (-\frac{(r-1)}{rc}\sigma^2)$$

$$\left.\left. + \sum_{ij}^{rc} \hat{\alpha}_i \hat{\beta}_j (\hat{\alpha}_i \hat{\beta}_j)\, \frac{\sigma^2}{rc}\right]\right\}$$

$$= E\left\{\frac{\sigma^2}{rc}\,(rc-c-r+1+c-1+r-1+1)\right\}$$

$$= E\left\{\sigma^2\right\} = \sigma^2$$

Thus, the variance of $\hat{\Upsilon}_1$ is σ^2.

It remains to demonstrate that $\hat{\Upsilon}_1$ has a normal distribution. Since our vector of parameter estimates $\underline{Z} = \underline{AX}$ has a normal distribution, the conditional distribution of the vector $\underline{Z}^{(2)} = (\hat{\epsilon}_{11}, \hat{\epsilon}_{12}, \cdots, \hat{\epsilon}_{r-1,c-1})$, given the vector of effect estimates $\underline{Z}^{(1)} = \underline{z}^{(1)}$, is normal [Anderson, 1958] with mean $\underline{\mu}^{(1)} + \underline{\Sigma}_{12}\underline{\Sigma}_{22}^{-1}(\underline{z}^{(2)} - \underline{\mu}^{(2)})$ and covariance matrix $\underline{\Sigma}_{11} - \underline{\Sigma}_{12}\underline{\Sigma}_{22}^{-1}\underline{\Sigma}_{21}$.

Since in this case $\underline{\Sigma}_{12} = \underline{\Sigma}_{21} = \underline{0}$, the mean is $\underline{\mu}^{(1)}$ and the covariance matrix is $\mathrm{Cov}(\hat{\underline{\epsilon}}_{ij}, \hat{\underline{\epsilon}}_{i'j'})$.

Thus, the conditional distribution of $\underline{Z}^{(2)}$, given $\underline{Z}^{(1)}$ is also the marginal distribution, since the vectors $Z^{(1)}$ and $Z^{(2)}$ are independent of each other.

Inasmuch as linear combinations of normal variates are again normally distributed, $\hat{\gamma}_1 \,\big|\, \hat{\alpha}_1, \cdots, \hat{\alpha}_{r-1}, \hat{\beta}_j, \cdots, \hat{\beta}_{c-1}$ has a conditional distribution which is normal with mean 0 and variance σ^2. As the conditional distribution is not dependent upon the conditions (the given parameter estimates), $\hat{\gamma}_1$ then has a normal distribution with mean 0 and variance σ^2.

Therefore $\hat{\gamma}_1^2 = \dfrac{(\Sigma \hat{\alpha}_i \hat{\beta}_j X_{ij})^2}{\Sigma \hat{\alpha}_i^2 \Sigma \hat{\beta}_j^2}$, which is Tukey's non-additivity sum of squares, has a σ^2 chi-squared distribution with 1 degree of freedom.

<u>Additional Individual Degrees of Freedom for Non-Additivity in the Two-way Design</u>

To approach the problem of the remaining one degree of freedom sums of squares, let us focus our attention on the vector $\underline{E}$ of rc error estimates. The whole set of error estimates have a joint multivariate normal distribution with mean vector 0 and singular covariance matrix $\underline{S}$. If we designate the (r-1)(c-1) by rc array of orthogonal contrasts used to partition the error sum of squares as matrix $\underline{D}$, $\underline{Y}=\underline{DE}$ is the vector of orthogonal contrasts. Now, since E is distributed according to $N(0,\underline{S})$, $\underline{Y}=\underline{DE}$ is distributed [Anderson, 1958] according to $N(0,\underline{DSD}')$. Now we shall show that $\underline{DSD}'$ is $\sigma^2\underline{I}$, with dimension (r-1)(c-1). We showed previously that $V\left(\dfrac{\sum\limits_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum\hat{\alpha}_i^2\sum\hat{\beta}_j^2}}\right) = \sigma^2$. This proof holds for any set of $\hat{\alpha}_i'$ and $\hat{\beta}_j'$ that are functions of the $\hat{\alpha}_i$ and $\hat{\beta}_j$, as long as $\sum\limits_{i}^{r}\hat{\alpha}_i' = 0 = \sum\limits_{j}^{c}\hat{\beta}_j'$. Thus, the main diagonal of $\underline{DSD}'$ consists of an array of σ^2.

To prove that the remainder of the matrix consists of zeros, we shall show that

$$\operatorname{Cov}\left(\frac{\sum\limits_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum\limits_{i}^{r}\hat{\alpha}_i^2\sum\limits_{j}^{c}\hat{\beta}_j^2}}\,,\;\frac{\sum\limits_{ij}^{rc}\hat{\alpha}_i'\hat{\beta}_j'\hat{\epsilon}_{ij}}{\sqrt{\sum\limits_{i}^{r}\hat{\alpha}_i'^2\sum\limits_{j}^{c}\hat{\beta}_j'^2}}\right) = 0$$

With the application of Lemma 3, the above expression becomes

$$E\left\{\operatorname{Cov}\left(\frac{\sum\limits_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum\limits_{i}^{r}\hat{\alpha}_i^2\sum\limits_{j}^{c}\hat{\beta}_j^2}}\,,\;\frac{\sum\limits_{ij}^{rc}\hat{\alpha}_i'\hat{\beta}_j'\hat{\epsilon}_{ij}}{\sqrt{\sum\limits_{i}^{r}\hat{\alpha}_i'^2\sum\limits_{j}^{c}\hat{\beta}_j'^2}}\;\middle|\;\hat{\alpha}_1,\cdots,\hat{\beta}_1,\cdots,\hat{\beta}_{c-1}\right)\right\}$$

$$+\operatorname{Cov}\left\{E\left(\frac{\sum\limits_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\epsilon}_{ij}}{\sqrt{\sum\limits_{i}^{r}\hat{\alpha}_i^2\sum\limits_{j}^{c}\hat{\beta}_j^2}}\;\middle|\;\hat{\alpha}_1,\cdots,\hat{\beta}_1,\cdots,\hat{\beta}_{c-1}\right),\right.$$

$$\left.E\left(\frac{\sum\limits_{ij}^{rc}\hat{\alpha}_i'\hat{\beta}_j'\hat{\epsilon}_{ij}}{\sqrt{\sum\limits_{i}^{r}\hat{\alpha}_i'^2\sum\limits_{j}^{c}\hat{\beta}_j'^2}}\;\middle|\;\hat{\alpha}_1,\cdots,\hat{\beta}_1,\cdots,\hat{\beta}_{c-1}\right)\right\}$$

The second term of this expression becomes $\operatorname{Cov}\{0,0\} = 0$.

The first expression may be expanded as

$$E\left\{\frac{1}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}\sqrt{\sum_i^r\hat{\alpha}_i'^2\sum_j^c\hat{\beta}_j'^2}}\left[\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\alpha}_i'\hat{\beta}_j'\operatorname{Var}(\hat{\epsilon}_{ij})\right.\right.$$

$$+\sum_{\substack{ijk\\ j\neq k}}^{r,c,c}\hat{\alpha}_i\hat{\alpha}_i'\hat{\beta}_j\hat{\beta}_k'\operatorname{Cov}(\hat{\epsilon}_{ij},\hat{\epsilon}_{ik})+\sum_{\substack{igj\\ i\neq g}}^{rrc}\hat{\alpha}_i\hat{\alpha}_g'\hat{\beta}_j\hat{\beta}_j'\operatorname{Cov}(\hat{\epsilon}_{ij},\hat{\epsilon}_{gj})$$

$$\left.\left.+\sum_{\substack{i,g,j,k\\ i\neq g\\ j\neq k}}^{rrcc}\hat{\alpha}_i\hat{\alpha}_g'\hat{\beta}_j\hat{\beta}_k'\operatorname{Cov}(\hat{\epsilon}_{ij},\hat{\epsilon}_{gk})\right]\right\}$$

$$=E\left\{\frac{1}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}\sqrt{\sum_i^r\hat{\alpha}_i'^2\sum_j^c\hat{\beta}_j'^2}}\left[\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\alpha}_i'\hat{\beta}_j'\frac{(r-1)(c-1)}{rc}\sigma^2\right.\right.$$

$$+\sum_{ij}\hat{\alpha}_i\hat{\alpha}_i'\hat{\beta}_j(-\hat{\beta}_j')\left(-\frac{(c-1)}{rc}\sigma^2\right)+\sum_{ij}\hat{\alpha}_i(-\hat{\alpha}_i')\hat{\beta}_j\hat{\beta}_j'\left(-\frac{(r-1)}{rc}\sigma^2\right)$$

$$\left.\left.+\sum_{ij}\hat{\alpha}_i\hat{\beta}_j(\hat{\alpha}_i'\hat{\beta}_j')\frac{\sigma^2}{rc}\right]\right\}$$

$$=E\left\{\frac{\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\alpha}_i'\hat{\beta}_j'\sigma^2}{\sqrt{\sum_i^r\hat{\alpha}_i^2\sum_j^c\hat{\beta}_j^2}\sqrt{\sum_i^r\hat{\alpha}_i'^2\sum_j^c\hat{\beta}_j'^2}}\right\}=E\left\{0\right\}=0$$

The above is equal to zero because $\sum_{ij}^{rc}\hat{\alpha}_i\hat{\beta}_j\hat{\alpha}_i'\hat{\beta}_j'\sigma^2=\sum_i^r\hat{\alpha}_i\hat{\alpha}_i'\sum_j^c\hat{\beta}_j\hat{\beta}_j'\sigma^2$ and the $\hat{\alpha}_i'$ and $\hat{\beta}_j'$ have been chosen so that $\sum_i^r\hat{\alpha}_i\hat{\alpha}_i'=0=\sum_j^c\hat{\beta}_j\hat{\beta}_j'$.

Extension to Higher-Way Designs

Tukey has indicated [1955] his general procedure for applying the single degree of freedom test of additivity to higher-way designs. The specific example illustrating his procedure is a 5 x 5 latin square.

Elston [1959], faced with a four-way classification with unequal subclass numbers, found it advisable to apply Tukey's test for non-additivity when interaction was found to be present---to see whether it indicated an appropriate transformation. However, he decided to apply the test on the basis of a two-way classification by collapsing three of the four classification variables to a single classification, for computational simplicity and because "it is much more likely that there exists a scale on which two factors are additive than one on which four factors are mutually additive".

The extension to higher-way designs which we should like to suggest does not have a single over-all degree of freedom for non-additivity corresponding to that of Tukey's; instead, there are separate degrees of freedom corresponding to the many various single degree of freedom sums of squares which could be isolated if the data were from a factorial experiment for which the levels are known.

For example, adding a third classification C at v levels, with effect estimates $\delta_1, \cdots, \delta_v$, would give us four separate single degree of freedom sums of squares of interest in testing for non-additivity. These would be:

$$\hat{\gamma}_1^2 = \frac{\left(\sum_{ijk}^{rcv} \hat{\alpha}_i \hat{\beta}_j \varepsilon_{ijk}\right)^2}{v \sum_i^r \hat{\alpha}_i^2 \sum_j^c \hat{\beta}_j^2}$$

$$\hat{\gamma}_2^2 = \frac{\left(\sum_{ijk}^{rcv} \hat{\alpha}_i \hat{\delta}_k \hat{\varepsilon}_{ijk}\right)^2}{c \sum_i^r \hat{\alpha}_i^2 \sum_k^v \hat{\delta}_k^2}$$

$$\hat{\gamma}_3^2 = \frac{\left(\sum_{ijk}^{rcv} \hat{\beta}_j \hat{\delta}_k \hat{\varepsilon}_{ijk}\right)^2}{r \sum_j^c \hat{\beta}_j^2 \sum_k^v \hat{\delta}_k^2}$$

$$\hat{\gamma}_4^2 = \frac{\left(\sum_{ijk}^{rcv} \hat{\alpha}_i \hat{\beta}_j \hat{\delta}_k \hat{\varepsilon}_{ijk}\right)^2}{\sum_i^r \hat{\alpha}_i^2 \sum_j^c \hat{\beta}_j^2 \sum_k^v \hat{\delta}_k^2}$$

The method suggested here is not applicable to the latin square design.

Literature Cited

1. Abraham, J. K., 1960, On an alternative method of computing Tukey's statistic for the latin square model, Biometrics 16:686-691, Note No. 154.

2. Anderson, T. W., 1958, Introduction to multivariate analysis, Wiley, New York.

3. Eisenhart, C., 1947, The assumptions underlying the analysis of variance, Biometrics 3:1-21.

4. Elston, Robert Claude, 1959, The estimation of genetic gain in milk yield due to sire selection over a period of time, Ph. D. thesis, Cornell University Library, September.

5. Elston, Robert Claude, 1961, On additivity in the analysis of variance, Biometrics 17:209-219.

6. Federer, Walter T., 1955, Experimental design, the MacMillan Company, New York.

7. Federer, Walter T., 1959, A note on additivity, Biometrics Unit, Cornell University, Mimeo. BU-74-M.

8. Hamaker, H. C., 1955, Experimental design in industry, Biometrics 11:257-286.

9. Harter, L. H. and Lum, Mary D., 1957, A note on Tukey's one degree of freedom for non-additivity, Paper presented at American Statistical Association meetings, Atlantic City, N. J., September 10-13.

10. Madow, William G., 1949, On the theory of systematic sampling, II, Ann. Math. Stat., 20:333-354.

11. Moore, Peter G. and Tukey, John W., 1954, Answer to query 112, Biometrics 10:562-568.

12. Robson, D. S., 1959, A simple method for constructing orthogonal polynomials when the independent variable is unequally spaced, Biometrics 15:187-191.

13. Robson, D. S. and Atkinson, G. F., 1960, Individual degrees of freedom for testing homogeneity of regression coefficients in a one-way analysis of covariance, Biometrics 16:593-605.

14. Snedecor, G. W., 1956, Statistical methods, Fifth Edition, Iowa State College Press, Ames, Iowa.

15. Tukey, John W., 1949, One degree of freedom for non-additivity, Biometrics 5:232-242.

16. Tukey, John W., 1955, Answer to query 113, Biometrics 11:111-113.

17. Ward, G. C. and Dick, I. D., 1952, Non-additivity in randomized block designs and balanced incomplete block designs, New Zealand Journal of Science and Technology 33:430-435.